

Taverna and myExperiment

Ravi Madduri

University of Chicago



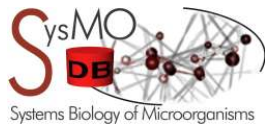
Accelerating Science




Taverna

Scientific workflow management system *for accessing public data services, assembling data processing and analysis pipelines and recording provenance.*

my experiment



BioCatalogue 

Social collaboration environments (“e-Laboratories”) *for sharing, curating and cataloguing personal, group and community contributed scientific assets.*

Taverna Capabilities

The screenshot displays the Taverna Workbench 2.1 interface, which is used for designing and executing workflow-based applications. The interface is divided into several main sections:

- Service panel:** Located on the left, it contains a list of available services and templates. The 'searchSimple' service is highlighted. Below the list, the 'Workflow explorer' and 'Details' tabs are visible. The 'Details' tab shows information for the 'program' input port, including its description, data format, example data, and computational type.
- Workflow diagram:** The central area shows a complex workflow diagram. It starts with 'Workflow input ports' (Sequence_id, Email_address) and flows through various service nodes (e.g., 'WSArrayData', 'InterProScan') to 'Workflow output ports' (InterProScan_XML_result, status, InterProScan_GFF, InterProScan_text_result, Job_ID).
- Service input port details:** The 'Details' tab for the 'program' input port provides the following information:
 - Description:** Specify blast type used: blastn, blastp, blastx, tblastn or tblastx.
 - Data format:** unknown
 - Example data:** blastp
 - Computational type:** xsd:string

Service Discovery

The screenshot displays the Taverna Workbench 2.1 interface. The main window shows a list of services (59/378) filtered by tag "ebi" and filtering criteria. The left sidebar contains a tree view of service categories and types. A dialog box titled "Filter Preview - untitled filter" is open, showing a preview of the filtered services. The right sidebar contains sections for Favourite Filters, Favourite Searches, and Search History.

Taverna Workbench 2.1

File Edit Insert View Workflows Advanced Help

Design Results myExperiment BioCatalogue

Search Filter Services About

Service Categories

- Sequence Analysis (97)
- Protein Sequence Analysis (74)
- Nucleotide Sequence Analysis (51)
 - Nucleotide Sequence Alignment (13)
 - Nucleotide Sequence Similarity (25)
 - Gene Prediction (10)
 - Promoter Prediction (1)
 - Transcription Factors (0)
 - Nucleotide Structure Prediction (0)
- Text Mining (12)
- Ontology (10)
- Phylogeny (2)
- Microarrays (39)
- Data Retrieval (58)
 - Sequence Retrieval (11)
 - Identifier Retrieval (12)
 - Pathway Retrieval (5)
 - Structure Retrieval (4)
 - Literature retrieval (11)
 - Image Retrieval (2)
- Genomics (11)
- Proteomics (27)
- Systems Biology (12)
- Biostatistics (13)
- Chemoinformatics (2)

Service Types

- SOAP (1141)
- REST (30)

Service Providers

- Submitters (Members)
- Submitters (Registries)
- Tags
- Tags (on Services)
- Tags (on Operations)
- Tags (on Inputs)

Showing top 17% of results of filtering services index by tag "ebi" and filtering criteria ...

Services (59/378)

- ChebiWebServiceService
- AccessionMapperService
- WSCitationImplService
- MiriamProviderService
- WSFastaService
- WSNCBIBlastService
- WSNCBIBlastService
- WSClustalW2Service
- WSKalignService
- WSMuscleService
- WSKalignService
- WSTCoffeeService
- Integr8ImplService
- whatizit
- WSDaliLiteService
- SBOPProviderService
- WSDBFetchServerLegacyService
- WSClustalWService
- WSEmbossService
- SBOPProviderService
- MiriamProviderService
- BioModelsWebServicesService
- picr
- picr

Filter Preview - untitled filter

Service Categories

- Sequence Analysis (97)
 - Nucleotide Sequence Analysis (51)
 - Gene Prediction (10)
- Data Retrieval (58)
 - Pathway Retrieval (5)

Service Types

- SOAP (1141)

Load OK

Favourite Filters

No favourite filters

Favourite Searches

- blast [services, service provid...]

Search History

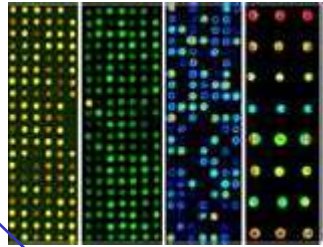
- ebi [all]
- blast [services, service provid...]

Filter More results All results

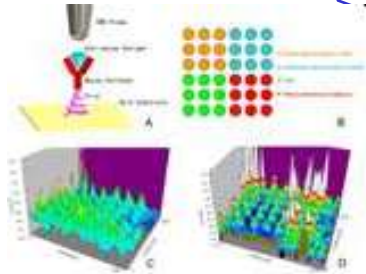
Lymphoma Prediction Workflow

caArray

MicroArray from
tumor tissue



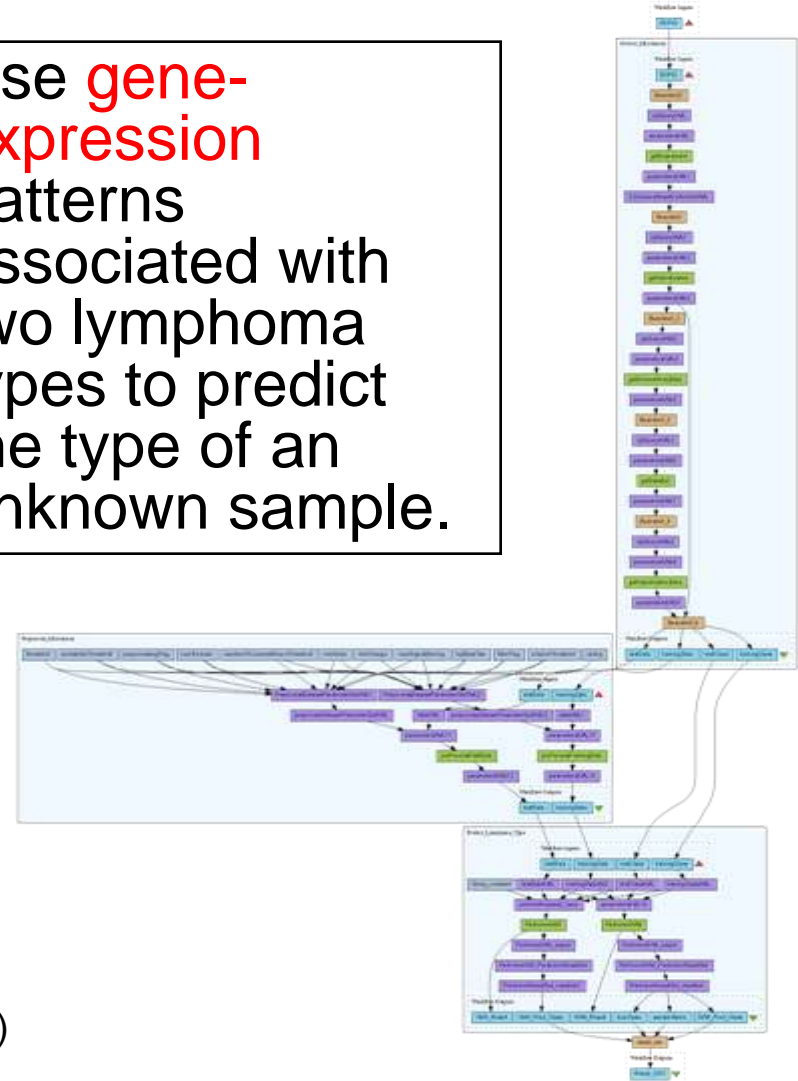
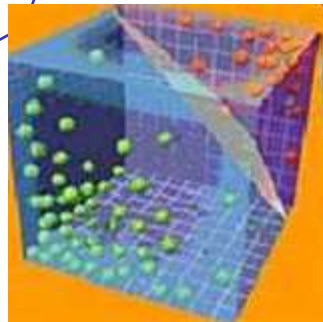
Microarray
preProcessing



Use **gene-expression** patterns associated with two lymphoma types to predict the type of an unknown sample.

Lymphoma
prediction

GenePattern

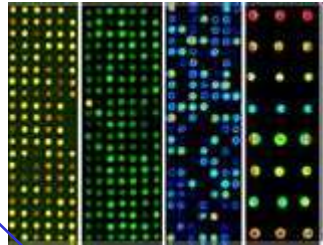


Ack. Juli Klemm, Xiaopeng Bian, Rashmi Srinivasa (NCI)
Jared Nedzel (MIT)

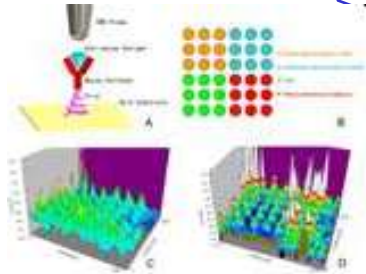
Lymphoma Prediction Workflow

caArray

MicroArray from
tumor tissue

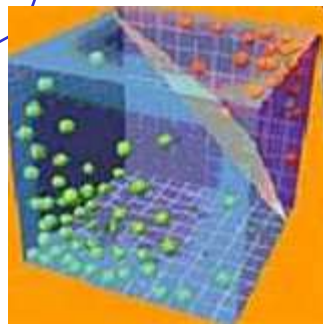


Microarray
preProcessing



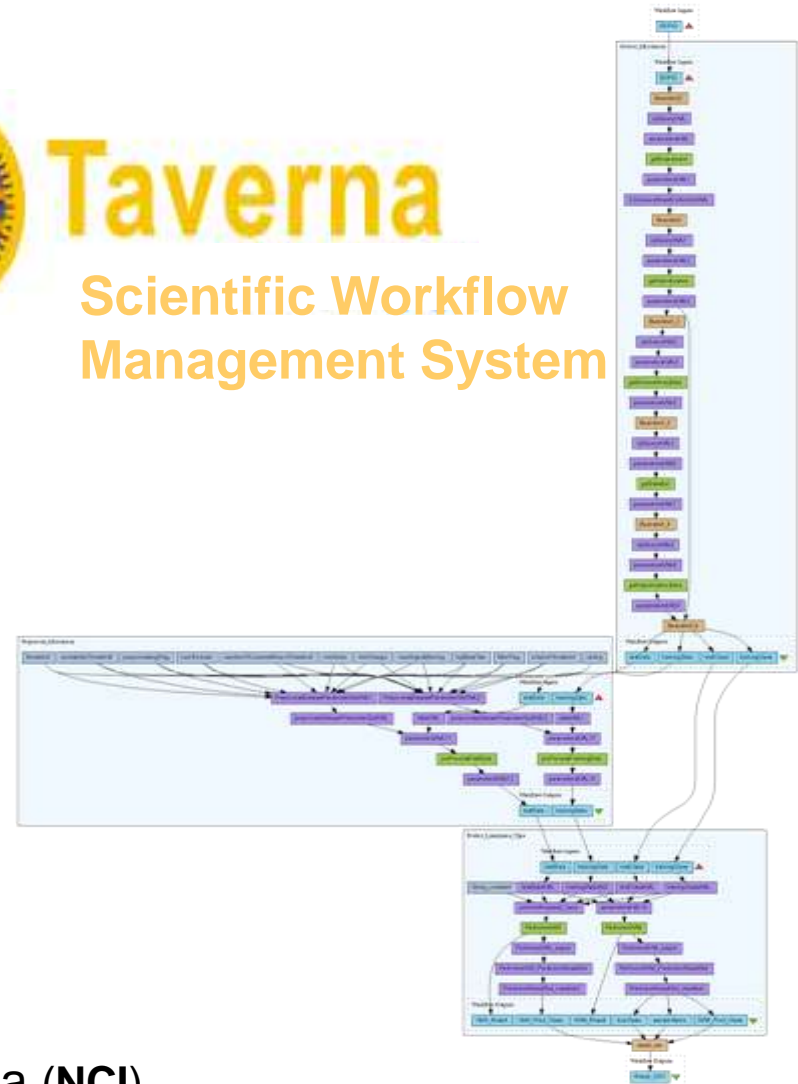
Lymphoma
prediction

GenePattern



Taverna

Scientific Workflow
Management System



Ack. Juli Klemm, Xiaopeng Bian, Rashmi Srinivasa (NCI)
Jared Nedzel (MIT)

Wei Tan Univ. Chicago

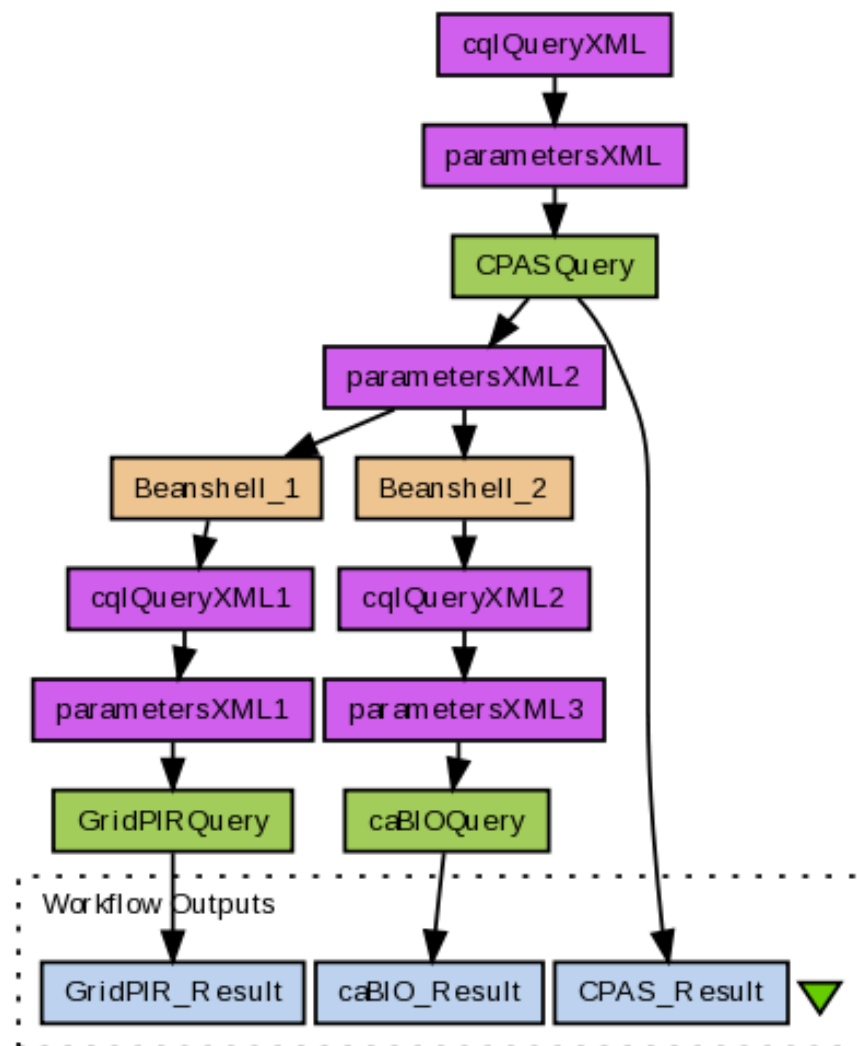
Protein sequence information query

- Scientific value

- To query protein sequence information out of 3 caGrid data services: **caBIO**, **CPAS** and **GridPIR**.
- To analyze a protein sequence from different data sources.

- Steps

- Querying **CPAS** and get the id, name, value of the sequence.
- Querying **caBIO** and **GridPIR** using the id or name obtained from **CPAS**.



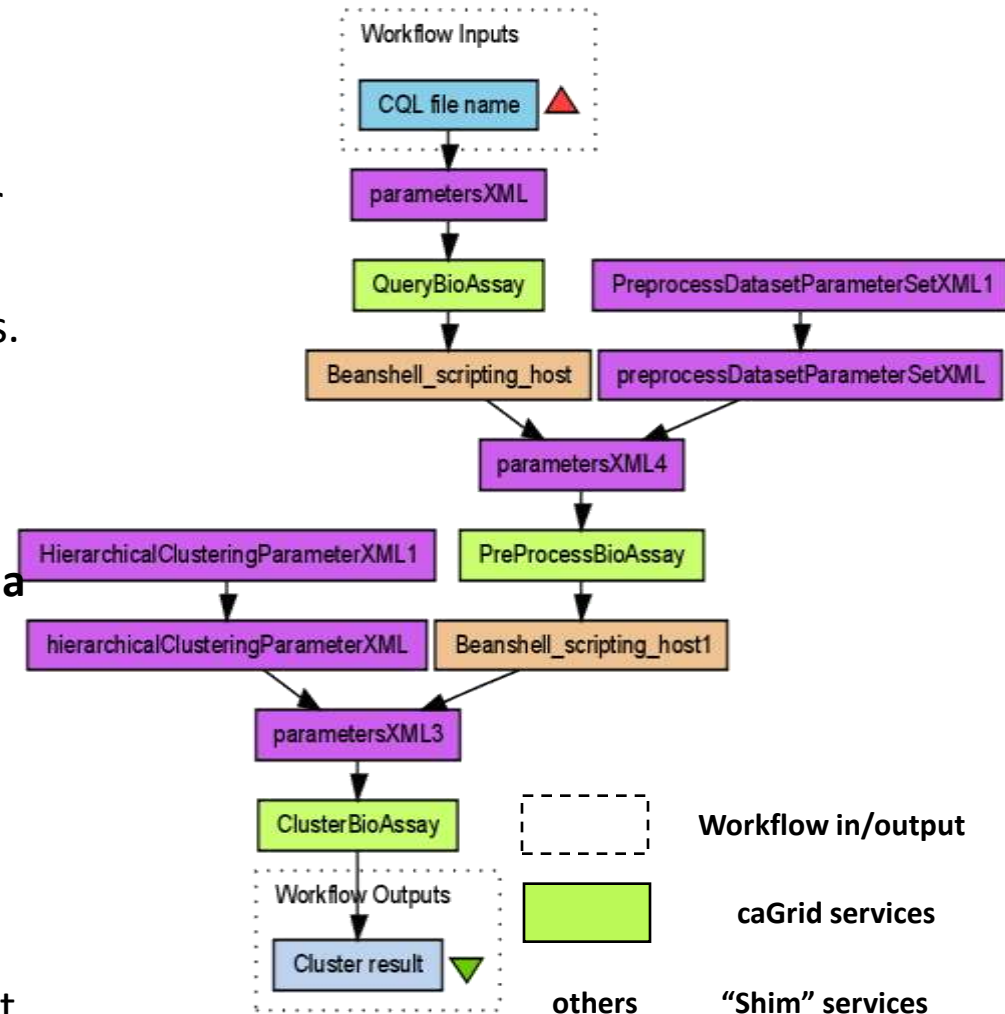
Microarray clustering

- Scientific value

- A common routine to group genes or experiments into clusters with similar profiles.
- To identify functional groups of genes.

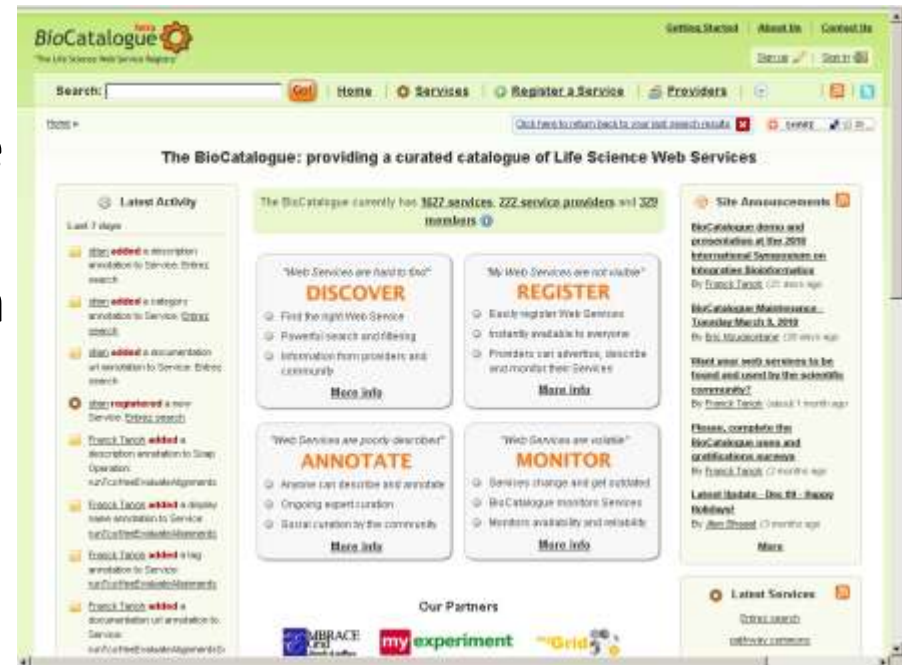
- Steps

- **Querying** and retrieving the microarray data of interest from a caArrayScrub **data** service at **Columbia University**
- **Preprocessing**, or normalize the microarray data using the GenePattern **analytical** service at the Broad Institute at **MIT**
- Running hierarchical **clustering** using the geWorkbench **analytical** service at **Columbia University**



BioCatalogue (www.biocatalogue.org)

- A “Web 2.0” catalogue for sharing, discovering and monitoring web services for the Life Sciences.
- Community and expert curation
- Community and provider contribution
- Launched mid 2009.
- Currently: 350+ members, 1600+ services, 110+ providers, 110+ different countries, 9000+ unique hits since launch



REST APIs
Linked Open Data
Software Open source BSD

my experiment

- “Facebook for Scientists” ...but different to Facebook!
- A repository of research methods
- A community social network of people and things
- A Social Virtual Research Environment
- Open source (BSD) Ruby on Rails app with REST and SPARQL interfaces
- Launched Nov 2007
- Basis or inspiration for multiple projects including BioCatalogue, MethodBox and SysmoDB

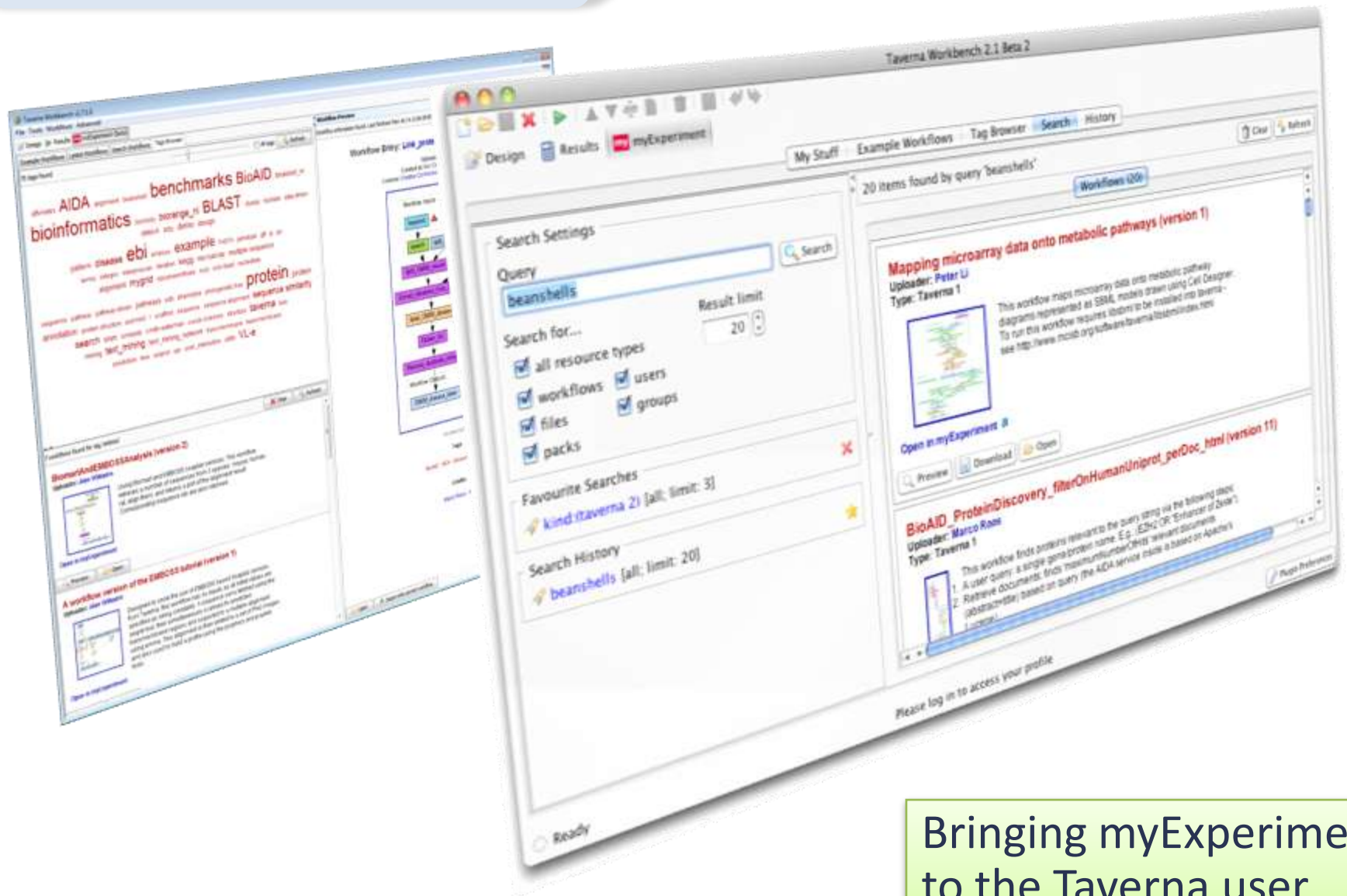
myExperiment currently has 3799 members, 228 groups, 1150 workflows, 325 files and 117 packs

myExperiment Features

- User Profiles
- Groups
- Friends
- Sharing
- Tags
- Workflows
- Developer interface
- Credits and Attributions
- Fine control over privacy
- Packs
- Federation
- Enactment



Taverna Plugins



Bringing myExperiment
to the Taverna user

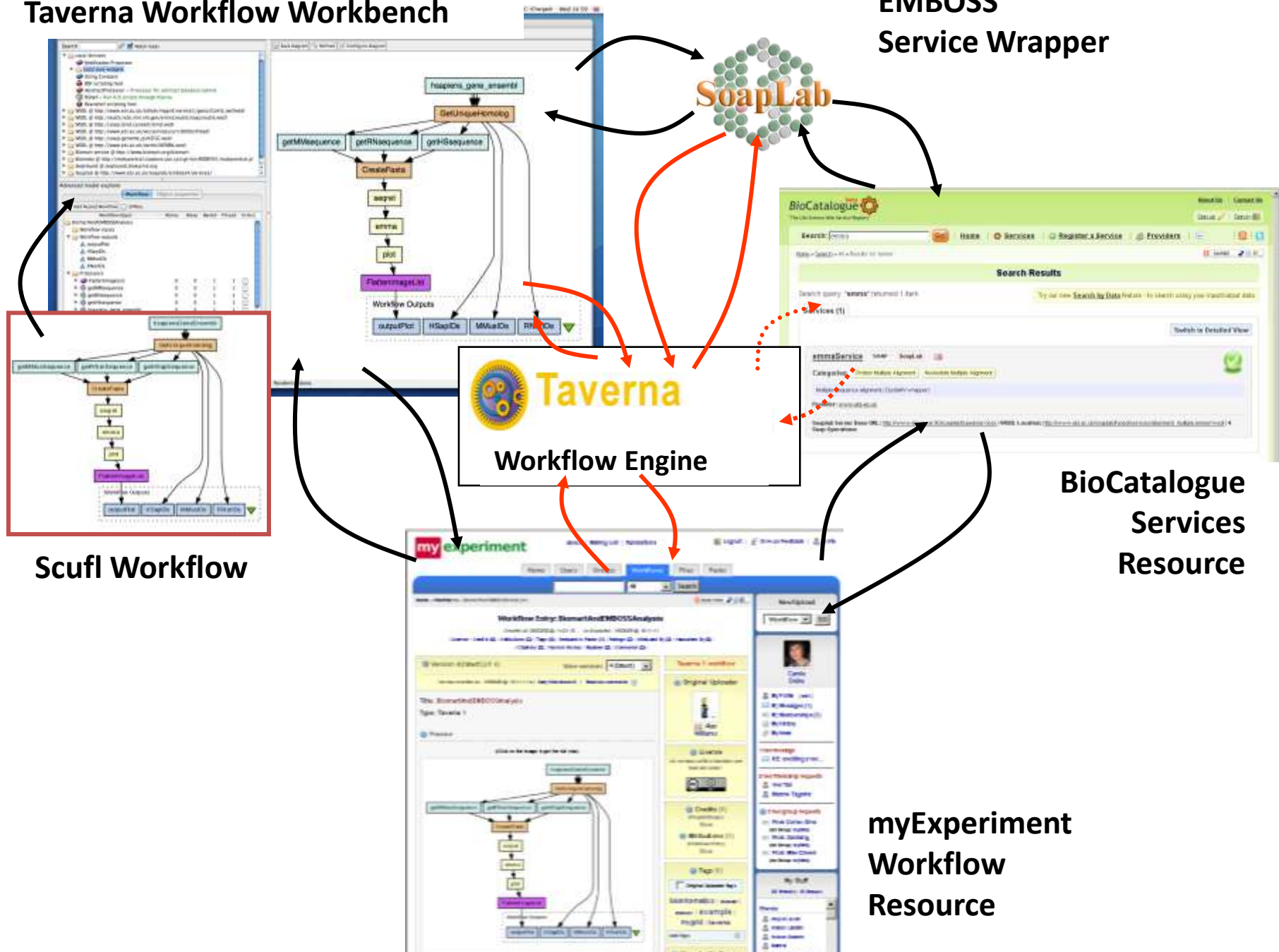
Google Gadgets



Bringing myExperiment
to the iGoogle user

Taverna Workflow Workbench

EMBOSS Service Wrapper



myExperiment - Search - Search

http://www.myexperiment.org/search?type=workflows&query=cabig

Google

Personal Dev Gmail Reader NYTimes Slashdot Rediff TOI Eenadu CNN Facebook Twitter CNET Wikipedia Subscribe... Read Later

myexperiment

About | Mailing List | Publications | Log in | Register | Give us Feedback | Invite

Home

Users

Groups

Workflows

Files

Packs

cabig

Workflows

Search

Home »

BOOK-MARK

Didn't find what you need? [Click here](#) to search external workflow repositories.

Search Results

14 Workflows found for "cabig".

Note: some items may not be visible to you, due to viewing permissions.


« previous


1

2


next »

Taverna 2

 **Lymphoma type prediction based on microarray data (v7)**




 View

Original Uploader

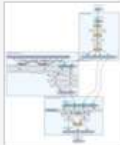


Wei Tan

Created: 11/05/10 @ 19:04:30 | Last updated: 11/05/10 @ 19:04:32

Credits:  Wei Tan  Madduri  Stian Soiland-Reyes

License: Creative Commons Attribution-Share Alike 3.0 Unported License



Scientific value


Using gene-expression patterns associated with DLBCL and FL to predict the lymphoma type of an unknown sample. Using SVM (Support Vector Machine) to classify data, and predicting the tumor types of unknown examples. Steps Querying training data from experiments stored in caArray. Preprocessing, or normalize the microarray data. Adding training and testing data into SVM service to get classification result.


Rating: 0.0 / 5 (0 ratings) | Versions: 7 | Reviews: 0 | Comments: 0 | Citations: 2

Viewed: 292 times | Downloaded: 120 times

Tags (11): cabig | cagrid | cancer research | classification | gene expression | globus | grid service | lymphoma | lymphoma workflow | microarray | support vector machine

Taverna 2

 **Support-Vector-Machine (SVM) based data classification (v2)**

 View

New/Upload

Workflow

Log in / Register

Username or Email:

Password:

Remember me: ☐

OR

Use OpenID:

(eg: name.myopenid.com)

Need an account?
[Click here to register](#)

Forgot Password?

Popular Tags

25 tags

[\[All Tags\]](#)

benchmarks

bioinformatics

example

localworker

mygrid

bio2rdf

BLAST

data integration

graph

index

ondex

d-grid

ebi

gws

kegg

microarray

ovtk

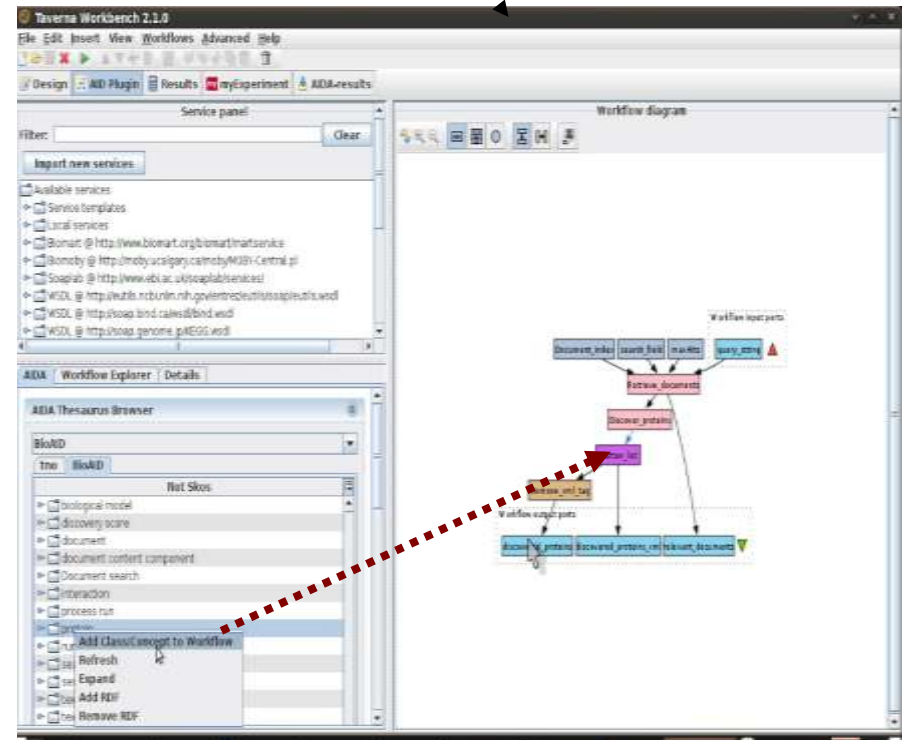
oxl

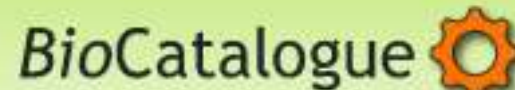
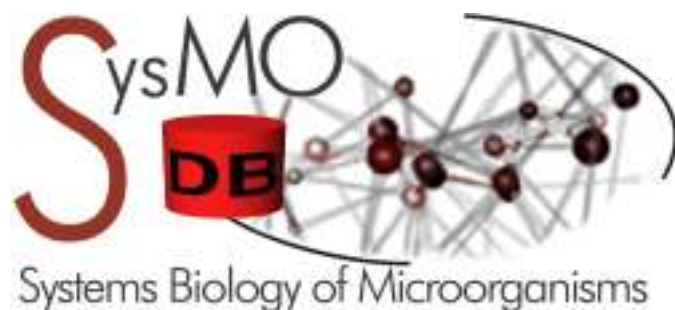
restful

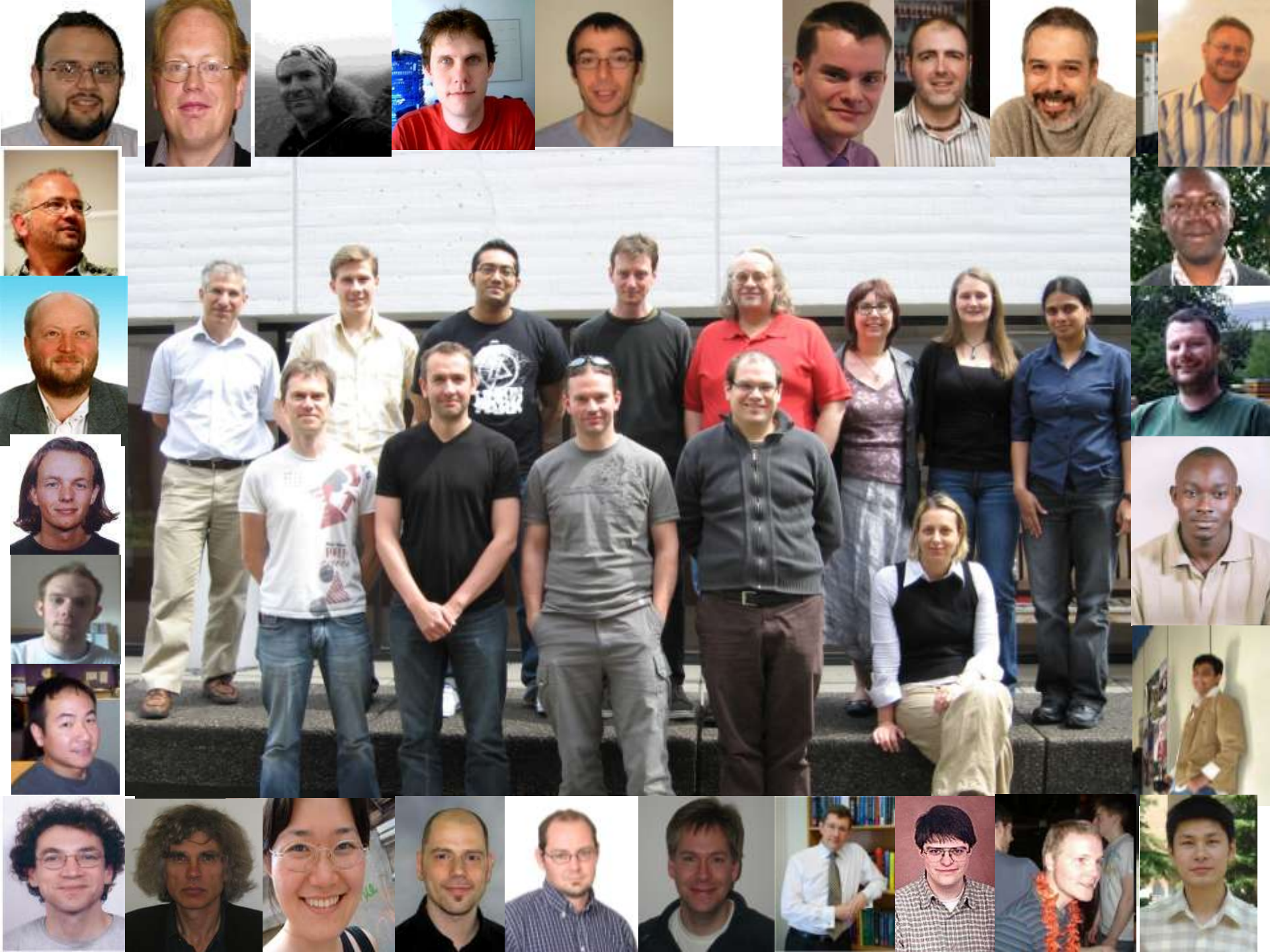
restlets

Taverna Labs

- Semantic Taverna
 - Semantic provenance
 - Open Provenance Model
 - Linked Open Data
 - Dutch NBIC Aida toolkit
 - Automated workflow planning through reasoning
 - e-Lico with U Zurich and Rapid-Miner
- Blogging the lab book
 - Blog3 with Southampton U







Contact

David De Roure

dder@ecs.soton.ac.uk

Carole Goble

carole.goble@manchester.ac.uk

Visit **myexperiment.org**

- Taverna: a tool for building and running workflows of services: Hull et al 2006, doi:10.1093/nar/gkl320
 - http://nar.oxfordjournals.org/cgi/content/abstract/34/suppl_2/W729
- myExperiment: a repository and social network for the sharing of bioinformatics workflows: Goble et al 2010, doi:10.1093/nar/gkq429
 - <http://nar.oxfordjournals.org/cgi/content/abstract/gkq429>
- BioCatalogue: a universal catalogue of web services for the life sciences Bhagat et al 2010, doi:10.1093/nar/gkq394
 - <http://nar.oxfordjournals.org/cgi/content/abstract/gkq394>